



## **Modelling Railway-Induced Passenger Delays in Multi-Modal Public Transport Networks: An Agent-Based Copenhagen Case Study Using Empirical Train Delay Data**

**Paulsen, Mads; Rasmussen, Thomas Kjær; Nielsen, Otto Anker**

*Publication date:*  
2018

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Paulsen, M., Rasmussen, T. K., & Nielsen, O. A. (2018). *Modelling Railway-Induced Passenger Delays in Multi-Modal Public Transport Networks: An Agent-Based Copenhagen Case Study Using Empirical Train Delay Data*. Paper presented at 14th Conference on Advanced Systems in Public Transport and TransitData 2018, Brisbane, Australia.

---

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Modelling Railway-Induced Passenger Delays in Multi-Modal Public Transport Networks

## An Agent-Based Copenhagen Case Study Using Empirical Train Delay Data

Mads Paulsen · Thomas Kjær Rasmussen · Otto Anker Nielsen

**Abstract** Due to lack of punctuality of public transport services, travel times of passengers are often uncertain. Whereas Automatic Vehicle Location (AVL) data makes it easy to measure the punctuality of public transport vehicles themselves, calculating door-to-door passenger delays is challenging as both the intended and realised routes of passengers have to be taken into account. This study introduces an agent-based MATSim simulation framework for evaluating passenger delays caused by delayed trains in multi-modal public transport systems. Three route choice strategies based on different levels of adaptiveness are considered, allowing passengers to intelligently deviate from their intended routes. Using empirical train delay data from the metropolitan area of Copenhagen for 65 weekdays in the autumn of 2014, the model concludes that the passenger delay distribution has a considerable higher standard deviation than the delay distribution of train arrivals. Additionally, the results reveal that a typical realised timetable would allow reduced overall passenger travel time compared to the published timetable.

**Keywords** Passenger Delays · Multi-Modal Public Transport Assignment · MATSim · Railway Delays

---

Mads Paulsen  
DTU Management Engineering  
Technical University of Denmark  
Bygningstorvet 116B  
2800 Kgs. Lyngby, Denmark  
Tel: +45 45 25 65 98  
E-mail: madsp@dtu.dk  
ORCID: 0000-0002-1445-0876

Thomas Kjær Rasmussen  
DTU Management Engineering  
Kgs. Lyngby, Denmark  
ORCID: 0000-0003-0979-9667

Otto Anker Nielsen  
DTU Management Engineering  
Kgs. Lyngby, Denmark  
ORCID: 0000-0001-7901-7021

---

## 1 Introduction

In most public transport networks a substantial share of passengers use more than one line to satisfy their transport needs. Due to the temporal dimension of public transport trips passenger delays are not additive over the trip segments, and the overall trip travel time will be unaffected by a delay on a trip segment that does not exceed the waiting time at the subsequent stop (Bates et al, 2001). On the contrary, a small vehicle delay of a suburban train connecting to a low-frequent local bus can cause severe passenger delays for all passengers having to use the bus. As such, determining passenger delays requires tracing every user through the public transport system while constantly ensuring that the correct departure is chosen conditioned on the delay of the previous segment. This makes them fundamentally different from car user delays which are found by summing the excess travel time of the links of the relevant path.

Paradoxically, public transport operators are generally evaluated on the basis of their vehicle punctuality and reliability, not the door-to-door punctuality of passengers (Noland and Polak, 2002). Although some researches have touched upon passenger perspectives in railway timetabling (Parbo et al, 2016), incorporating such passenger reliability measures in the actual evaluation of public transport systems is not done in practice. The first step to address this is to be able to calculate passenger consequences when public transport systems fail to meet full punctuality.

Ongoing research seems to focus on inferring such passenger delays by the use of data from individual passengers. Passenger delays were calculated in Jiang et al (2012); Sun et al (2016a,b) and Antos and Eichler (2016) for closed metro systems in Shanghai, Beijing, Shanghai and Washington D.C., respectively, using Automatic Fare Collection (AFC) data. Carrel et al (2015) combined smartphone data with AVL data to infer observed and intended trips through San Francisco's Muni network. Although the studies are capable of providing similar information about passenger delays as this study, the methods require data that typically cannot be assumed to be available across the entire network in general public transport systems. AFC data can only infer passenger delays in closed public transport systems, whereas tracking data from location-enabled devices such as smartphones is generally subject to a series of privacy issues (Rose, 2006; Gisdakis et al, 2015).

Instead, in this study we present an agent-based model framework for calculation of passenger delays by means of vehicle delays retrieved from AVL data generally collected in public transport systems. Passenger delays are evaluated by comparing the door-to-door travel times of intended routes (according to the planned timetable) to realised travel times modelled on the basis of empirical train delays. In the modelling of the latter, three route choice strategies with varying level of adaptiveness are analysed, making it possible for passengers to intelligently choose which vehicles to board on-the-go based on three different sets of assumptions. A detailed dataset consisting of train delays from 65 weekdays is used as input for the model, allowing aggregate, general analyses as well as disaggregate analyses of day-to-day variations in passenger delays. Such a model to infer passenger delays from vehicle delays has the potential to facilitate improvements of public transport systems towards better user experiences for the passengers using them.

---

The paper is organised as follows. Section 2 provides an introduction to existing literature on passenger delay modelling. The methodology behind the proposed model is presented in Section 3. Details about the case study is described in Section 4 with corresponding results reported in Section 5. The validity and perspectives of the results are discussed in Section 6, whereas Section 7 gives the conclusions of the study. Finally, Section 8 outlines directions of future work.

## 2 Literature Review

Determining passenger delays is a complex task as it requires knowledge about both the intended and realised route for every passenger in the system. In timetable-based public transport networks the planned routes and their associated travel times can be modelled in numerous ways (Liu et al, 2010). For instance by utilising a diachronic graph (Nuzzolo and Russo, 1994), a dual graph (Añez et al, 1996; Nielsen and Frederiksen, 2009) or a mixed line database (Hickman and Bernstein, 1997; Tong and Wong, 1999; Nielsen, 2000). However, routes found with the above methods are only valid when assuming full punctuality of services.

Unfortunately, full punctuality is rare, why dynamic adaptive route choice models are needed when modelling how passengers move through an actual unreliable schedule-based public transport network. Hickman and Bernstein (1997) introduced (and modelled on a toy network) four adaptive strategies ranging from allowing temporal choices while sticking to the same series of stop as intended (principle 1) to choosing an optimal path based on full a priori knowledge of the entire network (principle 4).

In the literature several of such adaptive route choice models for public transport have been proposed, both using frequency-based approaches (e.g. Gentile et al (2005); Teklu (2008)) and schedule-based approaches (e.g. Nuzzolo et al (2012)). Only a few of these deal with passenger delays explicitly. These will be presented in Section 2.1, before different kinds of models dealing with the issue less explicitly are presented in sections 2.2-2.4.

### 2.1 Explicit Passenger Delay Models

Nielsen et al (2001) and Nuzzolo et al (2001) both found passenger delays by averaging over several simulated days with vehicle delays drawn from statistical distributions. Both studies assumed full a priori knowledge of present and future delays was (principle 4, Hickman and Bernstein (1997)).

Poon et al (2004) presented a model with internally modelled queueing delays caused by passengers, but without including other causes of vehicle delays.

Landex and Nielsen (2006) used an approach with simulated train delays and temporal adaptive passenger route choices (principle 1, Hickman and Bernstein (1997)), but allowing full optimal adaptive route choice based on full information (principle 4, Hickman and Bernstein (1997)) after a passenger was delayed above a certain threshold. The model was applied to the suburban railway network of Copenhagen

---

using micro-simulated railway operations. Nielsen et al (2009) extended this by using realised delays provided by the railway operator.

Lijesen (2014) studied aspects of passenger delays using temporally adaptive passengers. Train delays were based on a distribution of train delays collected in the Netherlands for the second quarter of 2008. 16 direct city pairs were chosen for analysis. None of these routes required any transfers.

Zhu and Goverde (2017) introduced a model for a part of the Dutch train network used for evaluating passenger effects of adding additional trains under disruption. Although not the focus of their study, they explicitly suggested using the model for analysing passenger delays in future work.

In addition to these studies, the five studies with direct use of passenger data mentioned in the introduction (Jiang et al, 2012; Carrel et al, 2015; Sun et al, 2016a,b; Antos and Eichler, 2016) also dealt with passenger delays explicitly.

Furthermore, a series of studies have been somewhat associated with passenger delays without dealing with them explicitly. We have grouped these into three categories: studies only dealing with passengers delays caused by fail-to-board, studies only focusing on flow redistribution, and optimisation studies with the aim of improving operation once delays occur.

## 2.2 Capacity Caused Passenger Delay Models

A couple of studies deal with capacity-related delays in public transport systems, assuming that every vehicle runs on time but has finite capacity, possibly forcing some passengers to choose a different departure on-the-go.

Hamdouch and Lawphongpanich (2008) apply such strategies to a three-lines network while also considering departure time choice.

Nuzzolo et al (2012) presented an assignment model with full adaptive route choice within a predefined choice set and day-to-day departure time choice for the Fuorigrotta district of Naples. The analysis only focused on the number of fail-to-board instances.

Legara et al (2014) applied an agent-based model to a Singaporean network of 5 OD pairs validated by AFC data, and derived a critical capacity of the system.

## 2.3 Passenger Flow Redistribution Models

Some studies have had the aim of predicting passenger flows in case of emergencies or major disruptions, but without focusing on the time loss of passengers.

Li and Xu (2011) introduced a model for the Shanghai Rail Transit Network, where passenger flow is redistributed due to an artificial emergency situation.

Hong et al (2012) redistributed three classes of flows on a Shanghai mass transit network of 110 OD pairs based on a virtual emergency situation in the network. They calculated the total virtual loss (total passenger delay) for different durations of the emergency situation, but did not consider individual nor OD-specific passenger delays explicitly.

---

Xu et al (2014) considers different adaptive strategies including mode choice and route choice for passengers when encountering delays in a network of 240 OD pairs.

Similar adaptive strategies are used in Li and Zhu (2016) to make choices in artificially generated platform delays in the Shanghai metro network.

## 2.4 Passenger Delay Minimisation Models

Other studies deal with predefined initial railway delays, where the aim is to reschedule the trains in order to minimise passenger travel times.

Such studies include Dollevoet et al (2012) who applied passenger rerouting to a network consisting of 46 Dutch train stations, Sato et al (2013) who applied timetable rescheduling to a weekend timetable of the Chuo Line in Tokyo, Zhen and Jing (2016) who applied a train rescheduling model to a central part of the Beijing urban subway network, Corman et al (2017) who applied train rescheduling to a part the Dutch railway network in Amsterdam, Flevoland, and Utrecht, and finally Ghaemi et al (2018) who applied a disruption model to a railway network in the Dutch provinces Utrecht, Gelderland, and North Brabant.

None of the studies analyse individual or OD-specific passenger delays, but only with the total delay time which is used as a substantial part of the objective function.

## 3 Methodology

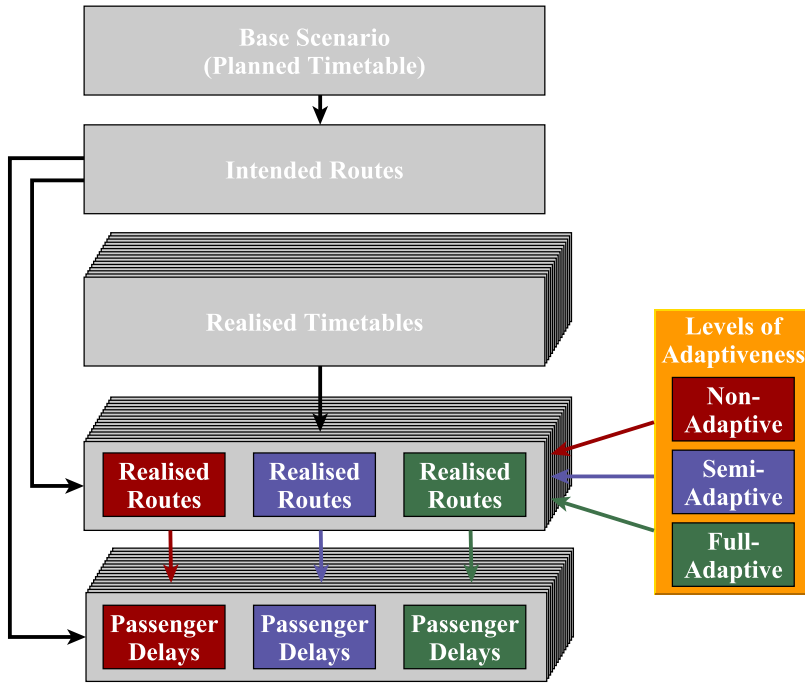
This study contributes to the literature by evaluating door-to-door passenger delays in a multi-modal public transport system of a large-scale metropolitan area using readily available train delay data. A diverse mix of transfers between low and high frequency lines for both buses and trains is secured through the multi-modality of the model, currently only represented in the literature when applying smartphone data of individual passengers (Carrel et al, 2015).

### 3.1 Framework

The model firstly identifies intended routes according to the timetable for all public transport trips, and then imposes the realised timetable under three different levels of passenger adaptiveness, see Section 3.2.

Such a process is repeated for each historical day with train delays available, allowing door-to-door travel patterns of each individual passenger to be modelled for each of these days. Comparing the trip travel times from these days to the base scenario where all trains run according to the timetable allows calculating passenger delays. The approach is illustrated in Figure 1, and formally written in Algorithm 1.

The model builds on version 0.9.0 of MATSim (Horni et al, 2016), an open-source transport simulator capable of modelling door-to-door transport on an individual level for both public transport and car traffic. The events-based public transport router extension of MATSim (Ordóñez, 2016), allows public transport users to reach optimal



**Fig. 1** The model framework consisting of a base scenario with a planned timetable providing intended routes, and scenarios for each historical day based on realised timetables resulting in realised routes.

intended routes through a day-to-day learning process by comparing the scores of performed routes across different iterations.

### 3.2 Levels of Adaptiveness

When the day is simulated, only adaptive choices in the time dimension (run choice) can be performed, in the sense that agents choose the first departing vehicle that can take them to the next stop in their planned route. This corresponds to principle 1 of Hickman and Bernstein (1997), and is denoted as the *non-adaptive* level of adaptiveness in this study.

In order to get a variety of plausible realised routes, two additional levels of adaptiveness are proposed – the *semi-adaptive* and the *full-adaptive*. In a traditional OD-based framework, the two strategies would collapse into a single strategy equivalent to principle 4 of Hickman and Bernstein (1997). However, in the door-to-door framework introduced for the current model the two strategies differ slightly.

The semi-adaptive strategy lets an agent walk to the first intended station/stop of the trip, at which point the agent seeks information about alternative routes – possibly departing from another stop/station, in which case the agent will walk to this stop/station.

---

**Algorithm 1**

---

```
1: Create the planned timetable,  $S^P$ .
2: Run MATSim with the events-based public transport router extension (Ordóñez,
   2016) using  $S^P$  to get the intended travel times,  $t_T^I$ , for all trips,  $T \in \mathcal{T}$ .
3: Create realised timetables,  $T_D^R$ , for all historical weekdays,  $D \in \mathcal{D}$ .
4: for all days,  $D \in \mathcal{D}$  do
5:   for all levels of adaptiveness,  $A \in \mathcal{A}$  do
6:     Simulate  $D$  in MATSim using the realised timetable,  $T_D^R$ , and the intended
       routes from 2, while allowing agents to make adaptive choices according
       to the level of adaptiveness,  $A$ , to obtain the corresponding realised travel
       times,  $t_T^{RA}$ , for all trips,  $T \in \mathcal{T}$ .
7:   for all trips,  $T \in \mathcal{T}$  do
8:     Find the corresponding passenger delay of  $T$ ,  $d_T^A$ , as the difference be-
       tween the intended travel time from 2 and realised travel time from 6,
       
$$d_T^A = t_T^{RA} - t_T^I. \quad (1)$$

9:   end for
10: end for
11: end for
```

---

In the full-adaptive strategy, the agent seeks information at the moment the public transport trip is initialised. This means that the agent can freely choose an initial station/stop to walk to in order to optimise the travel time.

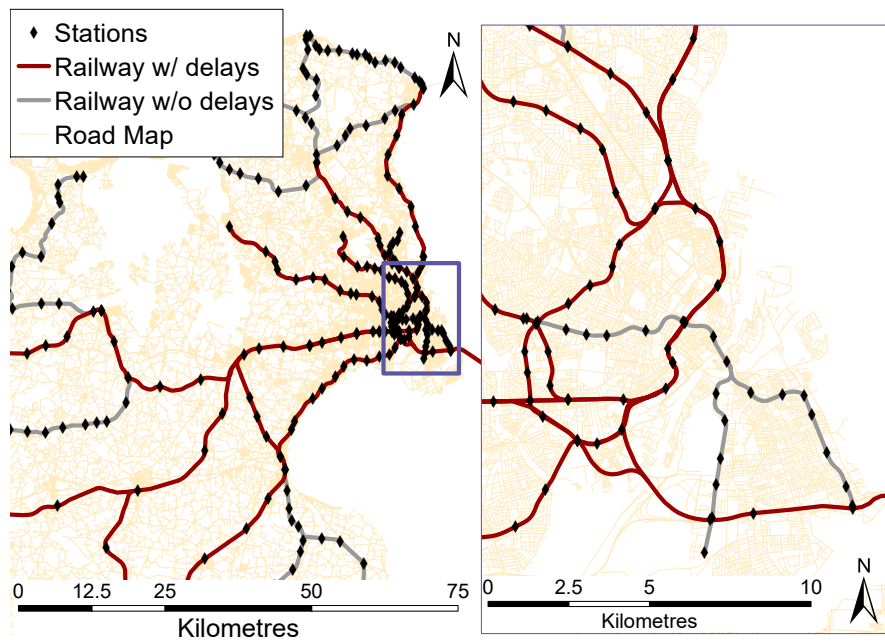
In both the semi- and full-adaptive strategies the agents are assumed to obtain full knowledge on past, current, and future delays when they seek information. Admittedly, this gives the agents abilities exceeding those of their real-life counterparts. Because this information is obtained at the beginning of each public transport trip in the full-adaptive strategy, the travel times from this should only be used as a lower bound estimate of passenger delays.

On the contrary, the non-adaptive strategy most likely forces the agents to be too conservative, why this should provide a reasonable estimate of the upper bound of the passenger delays.

As such, combining the obtained passenger delays from the non-adaptive and full-adaptive strategies will allow establishing an interval in which the real passenger delay is very likely to fall. The semi-adaptive approach will be placed somewhere in between the two but should not be considered as the true passenger delay, though, as the assumptions behind that strategy are debatable too.

Technically, the strategies are implemented by applying the events-based public transport router extension (Ordóñez, 2016) using the realised timetable and the original activity day plans for all agents for the adaptive strategy, and – for the semi-adaptive strategy – using modified day plans where the first transit walk segment of each public transport trip is fixed.





**Fig. 2** Map of the model area showing from which railway lines delay information was available.

#### 4 Case Study

The case study is based on a recently developed MATSim scenario for the metropolitan area of Copenhagen (Paulsen, 2016). This includes the entirety of the public transport system spread across different varieties of trains (24 lines), buses (400 lines), metro (2 lines) and ferries (1 line).

A base scenario using the planned timetable and 65 historical days from the autumn of 2014 with realised timetables have been run with the model. The planned timetable and the realised delays for the test period were provided by the railway manager Rail Net Denmark. The delay data covers all train runs in the region in the period excluding metro and local trains, see Fig. 2.

A 1% sample population was used in the tests presented in this paper. The agents that did not use the public transport system were filtered out, leaving the test dataset having 3,747 agents with a total of 7,889 daily public transport trips. 2,272 of these agents used the train system in at least one of the 65 days for at least one of the strategies. These agents providing 5,136 relevant public transport trips per simulated day have been used for the analysis presented in this paper.

The baseline was run using 200 iterations and a total running time of 13 hours and 44 minutes on a high performance computing node with two deca-core 2.60 GHz processors and 128 GB RAM available. The realised days ran with 75 iterations using the same hardware, and took approximately 4 hours per day per strategy. The realised days had to be run for both the semi-adaptive and the full-adaptive strategy, whereas

the non-adaptive strategy could be extracted from the initial iteration of either of the two strategies.

**Table 1** Proportion of passengers and trains arriving early, on time, and late for three different levels of adaptiveness.

	Non-Adaptive	Passengers Semi-Adaptive	Full-Adaptive	Trains
Delay $\leq$ -5 minutes [%]	2.1	6.3	8.2	0.4
Delay $\leq$ -1 minute [%]	8.4	13.6	15.7	11.9
Delay  $<$ 1 minute [%]	78.9	71.8	69.8	72.9
Delay $\geq$ 1 minute [%]	12.7	14.6	14.5	15.2
Delay $\geq$ 5 minutes [%]	5.3	6.9	6.7	3.0
Average Delay [minutes]	0.49	-0.43	-0.78	0.28
SD of Delay [minutes]	5.80	19.18	20.06	2.82
Number of Observations	333,804	333,757	333,686	2,362,880

## 5 Results

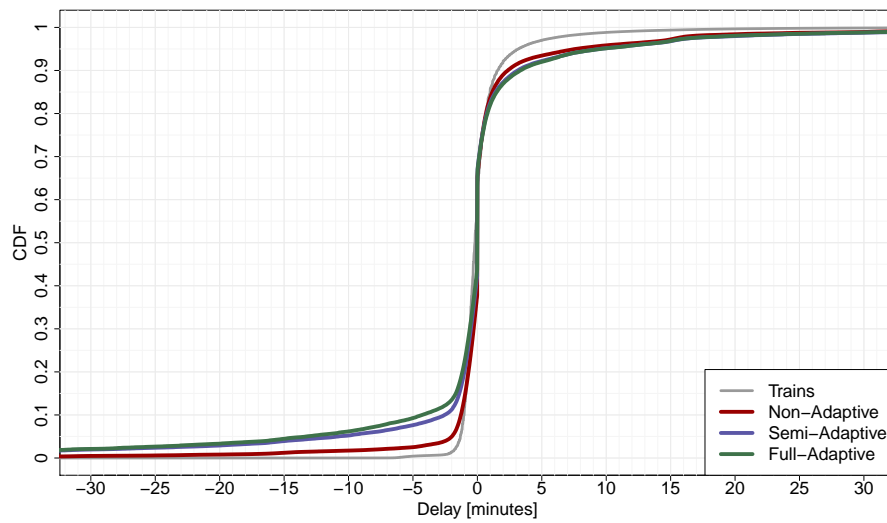
The aggregate results shown in Table 1 illustrate that the average trip delay for agents using the railway system is roughly zero for all of the three considered strategies. The average delay is positive for the non-adaptive strategy, whereas the other two strategies have a negative average delay. They also have a much larger standard deviation (19.18 – 20.06 minutes) than that of the non-adaptive strategy (5.80 minutes), although the trips regardless of strategy in general have a higher standard deviation than the trains (2.82 minutes).

It is seen that the majority of passengers (69.8-78.9%) arrive within one minute of their intended arrival time, regardless of choice strategy. It is also seen that all three strategies have a proportion of passengers being delayed more than one minute (12.7-14.6%) that is lower than that of trains (15.2%). However, with a threshold of 5 minutes, all three strategies suggest it more likely for passengers to be delayed than for single trains, with 5.3-6.9% being delayed as opposed to 3.0% of trains.

The greatest difference between the levels of adaptiveness is seen regarding time savings (negative delays). Whereas 8.4% and 2.1% of passengers save more than 1 and 5 minutes, respectively, using the non-adaptive strategy, the numbers are significantly higher using the semi- and full-adaptive strategies, with the former having 13.6% and 6.3% and the latter having 15.7% and 8.2% arriving at least 1 and 5 minutes before, respectively.

As expected, it is seen that train punctuality is considerably higher than the passenger punctuality. This is even more evident in Figure 3 where the empirical cumulative distribution functions of the passenger and train delays of the current study are illustrated.

Fig. 4 shows the survival functions for four distributions of savings and four distributions of delays. It is interpreted such that the value on the second axis is the



**Fig. 3** Empirical cumulative distribution functions delays of train arrival at each station and passenger trip delays for three levels of adaptiveness.

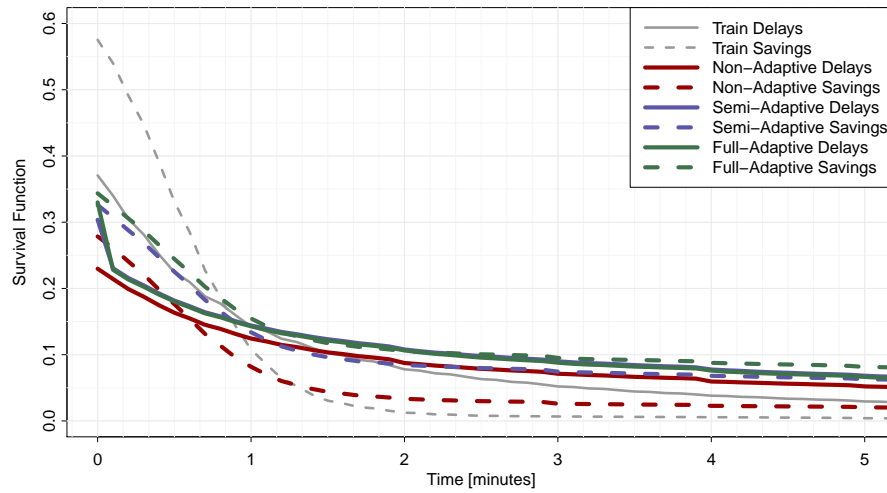
proportion of relevant trips having a delay/saving larger than or equal to the number on the first axis.

It is seen that it is very common for trains to arrive a little early, but that delays are more common than savings when we exceed 50 seconds. The same tendency holds for the non-adaptive and the semi-adaptive strategy, where the break-even points are found at 35 seconds and 50 seconds, respectively. This is most likely caused by trips having a train segment as the last segment of the trip, as trains commonly arrive a little early at the stations – 15.9% of trains in the dataset arrive between 1 and 35 seconds earlier than scheduled.

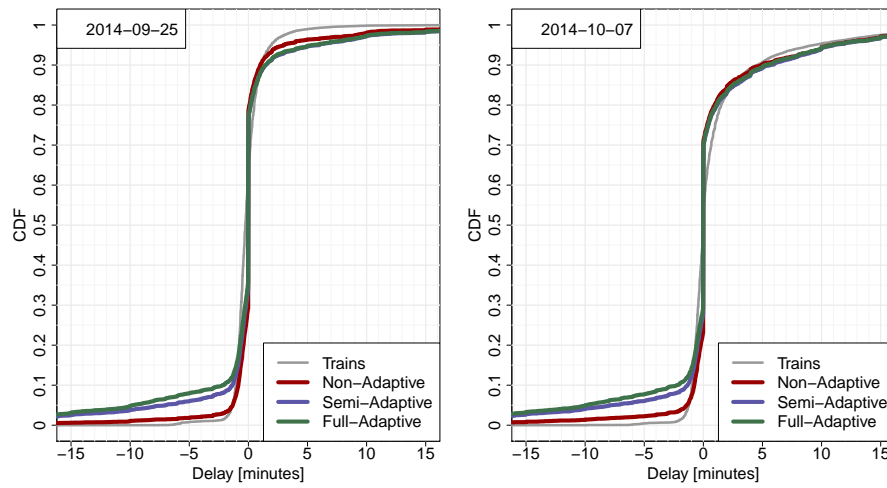
For the full-adaptive strategy there seems to be a break-even point at 70 seconds, but that the savings and delays are equally likely from here and up to about two and a half minute. Exceeding this point savings are again more likely than delays, which also continues to be the case for delays/savings exceeding five minutes, with 3.3% of passengers saving more than 15 minutes, whereas only 2.0% of passengers are 15 minutes delayed.

It is also worth noting that the survival functions for semi- and full-adaptive delays are almost indistinguishable, and that the survival function for the non-adaptive delays is also rather close to this value. This could indicate, that although great improvements can be made for some agents when using a high level of adaptiveness, for some agents there are no good alternatives, why a delay in such case is inevitable.

Finally, it is worth mentioning that although some variation between days does exist (see Fig. 5), the overall conceptual tendencies remains the same for all of the days in the test period.



**Fig. 4** Empirical survival functions of delays and time savings for train arrivals at each station and passenger trips for three levels of adaptiveness.



**Fig. 5** Cumulative distribution functions of delay for the least (left) and most (right) delayed day of the test period.

## 6 Discussion

The proportion of passengers arriving late is in line with findings in Nielsen et al (2009) that showed that 15% of passengers arrive late when only considering the sub-urban railway network of Copenhagen. Likewise, the proportion of passengers with

---

time saving above one minute was in line with the adaptive strategies of this study (17.3-19.6%). However, the average delay was much larger in Nielsen and Frederiksen (2009), reaching 7.5-8.4 minutes as opposed to -0.78-0.49 in this study. The fact that many passengers have additional connections after using the suburban railway network may possibly explain the difference in average delay between the two studies.

The attentive reader might have noticed that all routes available for the non-adaptive strategy are also available using the semi- and full-adaptive strategies. Likewise any semi-adaptive route is also a feasible full-adaptive route. Due to this, it seems counter-intuitive that larger delays are more common using the semi- and full-adaptive strategies than using the non-adaptive strategy. However, some reasonable explanations for this do exist.

Firstly, all agents are optimising their daily plan according to a generalised travel time function, where waiting time is penalised 60% harder than in-vehicle and walking travel time. Thus, a trip may have a lower generalised travel time than another trip even though the actual travel time is greater.

Secondly, the agents optimise entire day plans. If all activities are sufficiently long (or delays sufficiently short), then the end time of all activities remain the same regardless of the used routes. However, if an activity is relatively short and the delay is relatively large, the agent might arrive at an activity after it is scheduled to terminate. An agent will in this case continue towards the next activity of his/her plan from this moment, causing the departure time of the trip to be dependent on the route of subsequent trips. In such cases the trips are incomparable why shorter travel times may occur in such cases.

On the other end of the distribution, it might seem off that both the semi-adaptive and full-adaptive strategy generally reveal shorter travel time than the planned travel times, existing literature points towards it being plausible.

Parbo et al (2014) showed with a large-scale bi-level optimisation algorithm that the overall waiting time in the public transport system in Denmark can be reduced by more than 5% without influencing in-vehicle and walking travel time considerably.

Similarly, Fonseca et al (2018) showed through a bi-level optimisation of strategic buses in the metropolitan area of Copenhagen, that the excess transfer time between such buses and trains could be reduced by almost 2 minutes on average

The two studies reveal that the actual timetable for the given time was indeed sub-optimal. This suggests that it is indeed possible that the realised operations – which basically can be seen as the planned operations with some random mutations – could lead to a more well-integrated public transport system than the planned operations.

## 7 Conclusions

This paper has presented an agent-based framework for determination of passenger delays in large-scale multi-modal transport systems based on AVL data from trains. In the application to the metropolitan area of Copenhagen based on 65 realised timetables, the model required no additional empirical passenger information and was still

---

able to calculate door-to-door delays of individual agents across different modes of the public transport system for three different levels of adaptiveness.

Results showed that the passengers delays depend a lot on the adaptiveness of the route choice strategy. However, it was also seen that passenger travel time in general is more volatile than train travel times, and that the volatility increases with the level of adaptiveness. The semi-adaptive and – in particular – the full-adaptive strategy commonly allowed relatively large travel time savings, forcing the average passenger delay to be negative. The findings show that the planned timetable used during the test period has room for improvement, as the realised timetables generally allowed for shorter travel times than the planned one.

## 8 Future Work

Future work includes strengthening the multi-modal aspect of the model also including AVL data from buses in the realised timetables. Another extension would be to increase the population sample to 10% to provide a better coverage of the model area. This would allow carrying out detailed analyses such as investigating the effect of transfer types (to/from bus/train) and stations, which currently would not be sufficiently comprehensive due to lack of data. Such analysis could be supplemented by actual geostatistical analysis such as kriging. This would reveal spatial interdependencies and identify the most delay-prone transport hubs in the network.

Methodologically, an obvious next step includes implementing the adaptive route strategies as actual within-day replanning (Dobler and Nagel, 2016) in MATSim as currently available for car traffic, e.g. in Kaddoura and Nagel (2018). This would greatly speed up the computation time whilst being a more intuitive way to approach the problem at hand.

**Acknowledgements** This study is a part of the IPTOP (Integrated Public Transport Optimisation and Planning) research project funded by Innovation Fund Denmark, which the authors would like to thank for their financial support.

## References

- Añez J, De La Barra T, Pérez B (1996) Dual graph representation of transport networks. *Transportation Research Part B: Methodological* 30(3):209–216, DOI 10.1016/0191-2615(95)00024-0, URL <http://linkinghub.elsevier.com/retrieve/pii/0191261595000240>
- Antos J, Eichler MD (2016) Tapping into Delay: Assessing Rail Transit Passenger Delay With Data from a Tap-In, Tap-Out Fare System. *Transportation Research Record: Journal of the Transportation Research Board* 2540(May):76–83, DOI 10.3141/2540-09, URL <http://trrjournalonline.trb.org/doi/10.3141/2540-09>
- Bates J, Polak J, Jones P, Cook A (2001) The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review* 37(2-

- 
- 3):191–229, DOI 10.1016/S1366-5545(00)00011-9, URL <http://linkinghub.elsevier.com/retrieve/pii/S1366554500000119>
- Carrel A, Lau PS, Mishalani RG, Sengupta R, Walker JL (2015) Quantifying transit travel experiences from the users' perspective with high-resolution smartphone and vehicle location data: Methodologies, validation, and example analyses. *Transportation Research Part C: Emerging Technologies* 58:224–239, DOI 10.1016/j.trc.2015.03.021, URL <http://linkinghub.elsevier.com/retrieve/pii/S0968090X15001060>
- Corman F, D'Ariano A, Marra AD, Pacciarelli D, Samà M (2017) Integrating train scheduling and delay management in real-time railway traffic control. *Transportation Research Part E: Logistics and Transportation Review* 105:213–239, DOI 10.1016/j.tre.2016.04.007, URL <http://dx.doi.org/10.1016/j.tre.2016.04.007>
- Dobler C, Nagel K (2016) Within-Day Replanning. In: Horni A, Nagel K, Axhausen KW (eds) *The Multi-Agent Transport Simulation MATSim*, Ubiquity Press, London, chap 30, pp 196–211, DOI 10.5334/baw
- Dollevoet T, Huisman D, Schmidt M, Schöbel A (2012) Delay Management with Rerouting of Passengers. *Transportation Science* 46(1), DOI 10.1287/trsc.1110.0375, URL <https://doi.org/10.1287/trsc.1110.0375>
- Fonseca JP, van der Hurk E, Roberti R, Larsen A (2018) A matheuristic for transfer synchronization through integrated timetabling and vehicle scheduling. *Transportation Research Part B: Methodological* 109:128–149, DOI 10.1016/j.trb.2018.01.012, URL <http://linkinghub.elsevier.com/retrieve/pii/S0191261517304915>
- Gentile G, Nguyen S, Pallottino S (2005) Route Choice on Transit Networks with Online Information at Stops. *Transportation Science* 39(3):289–297, DOI 10.1287/trsc.1040.0109, URL <http://pubsonline.informs.org/doi/abs/10.1287/trsc.1040.0109>
- Ghaemi N, Zilko AA, Yan F, Cats O, Kurowicka D, Goverde RM (2018) Impact of railway disruption predictions and rescheduling on passenger delays. *Journal of Rail Transport Planning & Management* DOI 10.1016/j.jrtpm.2018.02.002, URL <http://linkinghub.elsevier.com/retrieve/pii/S2210970617300872>
- Gisdakis S, Manolopoulos V, Tao S, Rusu A, Papadimitratos P (2015) Secure and Privacy-Preserving Smartphone-Based Traffic Information Systems. *IEEE Transactions on Intelligent Transportation Systems* 16(3):1428–1438, DOI 10.1109/TITS.2014.2369574, URL <http://ieeexplore.ieee.org/document/6991589/>
- Hamdouch Y, Lawphongpanich S (2008) Schedule-based transit assignment model with travel strategies and capacity constraints. *Transportation Research Part B: Methodological* 42(7-8):663–684, DOI 10.1016/J.TRB.2007.11.005, URL <https://www.sciencedirect.com/science/article/pii/S0191261507001361?via%3Dihub>
- Hickman MD, Bernstein DH (1997) Transit Service and Path Choice Models in Stochastic and Time-Dependent Networks. *Transportation Science* 31(2):129–146, DOI 10.1287/trsc.31.2.129, URL <http://pubsonline.informs.org/doi/abs/10.1287/trsc.31.2.129>

- 
- Hong L, Gao J, Xu RH (2012) Influence of Emergency Passenger Flow Distribution in Urban Rail Network. *Advanced Materials Research* 450-451:295–301, DOI 10.4028/www.scientific.net/AMR.450-451.295, URL <https://www.scientific.net/AMR.450-451.295>
- Horni A, Nagel K, Axhausen KW (eds) (2016) *The Multi-Agent Transport Simulation MATSim*. Ubiquity Press, London, DOI 10.5334/baw, URL <https://www.ubiquitypress.com/site/books/10.5334/baw/>
- Jiang ZB, Li F, Xu RH, Gao P (2012) A simulation model for estimating train and passenger delays in large-scale rail transit networks. *Journal of Central South University* 19(12):3603–3613, DOI 10.1007/s11771-012-1448-9, URL <http://link.springer.com/10.1007/s11771-012-1448-9>
- Kaddoura I, Nagel K (2018) Using real-world traffic incident data in transport modeling. *Procedia Computer Science* 130:880–885, DOI 10.1016/J.PROCS.2018.04.084, URL <https://www.sciencedirect.com/science/article/pii/S1877050918304460>
- Landex A, Nielsen OA (2006) Simulation of disturbances and modelling of expected train passenger delays. In: *Computers in Railways X*, WIT Press, Southampton, UK, WIT Transactions on The Built Environment, Vol 88, vol 1, pp 521–529, DOI 10.2495/CR060521, URL <http://library.witpress.com/viewpaper.asp?pcode=CR06-052-1>
- Legara EF, Monterola C, Lee KK, Hung GG (2014) Critical capacity, travel time delays and travel time distribution of rapid mass transit systems. *Physica A: Statistical Mechanics and its Applications* 406:100–106, DOI 10.1016/j.physa.2014.02.058, URL <http://linkinghub.elsevier.com/retrieve/pii/S0378437114001678>
- Li F, Xu R (2011) Simulation of passenger distribution in urban rail transit in case of emergencies. In: *Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE)*, IEEE, vol 3, pp 1933–1936, DOI 10.1109/TMEE.2011.6199592, URL <http://ieeexplore.ieee.org/document/6199592/>
- Li W, Zhu W (2016) A dynamic simulation model of passenger flow distribution on schedule-based rail transit networks with train delays. *Journal of Traffic and Transportation Engineering (English Edition)* 3(4):364–373, DOI 10.1016/j.jtte.2015.09.009, URL <http://linkinghub.elsevier.com/retrieve/pii/S2095756415305833>
- Lijesen MG (2014) Optimal Traveler Responses to Stochastic Delays in Public Transport. *Transportation Science* 48(2):256–264, DOI 10.1287/trsc.2013.0465, URL <http://pubsonline.informs.org/doi/abs/10.1287/trsc.2013.0465>
- Liu Y, Bunker J, Ferreira L (2010) Transit Users' Route-Choice Modelling in Transit Assignment: A Review. *Transport Reviews* 30(6):753–769, DOI 10.1080/01441641003744261, URL <http://www.tandfonline.com/doi/abs/10.1080/01441641003744261>
- Nielsen OA (2000) A stochastic transit assignment model considering differences in passengers utility functions. *Transport Research Part B* 34:377–402
- Nielsen OA, Frederiksen RD (2009) Large-scale schedule-based transit assignment - further optimization of the solution algorithms. *Operations Research/ Computer*



- 
- Science Interfaces Series 46:119–144, URL <http://www.scopus.com/inward/record.url?eid=2-s2.0-84867783635&partnerID=tZ0tx3y1>
- Nielsen OA, Overgård CH, Daly A (2001) A Large-scale model system for the Copenhagen-Ringsted railway project. In: Hensher D (ed) *Travel Behaviour Research*, 1st edn, Elsevier, Oxford, UK, chap 35, pp 603–626, URL <http://www.forskningsdatabasen.dk/en/catalog/2389465253>
- Nielsen OA, Landex O, Frederiksen RD (2009) Passenger delay models for rail networks. In: *Schedule-Based Modeling of Transportation Networks*, Springer US, Boston, MA, chap 2, pp 27–49, DOI 10.1007/978-0-387-84812-9\_2, URL [http://link.springer.com/10.1007/978-0-387-84812-9\\_2](http://link.springer.com/10.1007/978-0-387-84812-9_2)
- Noland RB, Polak JW (2002) Travel time variability: A review of theoretical and empirical issues. *Transport Reviews* 22(1):39–54, DOI 10.1080/01441640010022456, URL <http://www.tandfonline.com/doi/abs/10.1080/01441640010022456>
- Nuzzolo A, Russo F (1994) An equilibrium assignment model for intercity transit networks. In: *Proceedings of TRISTAN II Conference*, Capri, Italy
- Nuzzolo A, Russo F, Crisalli U (2001) A Doubly Dynamic Schedule-based Assignment Model for Transit Networks. *Transportation Science* 35(3):268–285, DOI 10.1287/trsc.35.3.268.10149, URL <http://pubsonline.informs.org/doi/abs/10.1287/trsc.35.3.268.10149>
- Nuzzolo A, Crisalli U, Rosati L (2012) A schedule-based assignment model with explicit capacity constraints for congested transit networks. *Transportation Research Part C: Emerging Technologies* 20(1):16–33, DOI 10.1016/j.trc.2011.02.007, URL <http://linkinghub.elsevier.com/retrieve/pii/S0968090X11000337>
- Ordóñez SA (2016) New Dynamic Events-Based Public Transport Router. In: *The Multi-Agent Transport Simulation MATSim*, Ubiquity Press, London:, chap 19, pp 126–136
- Parbo J, Nielsen OA, Prato CG (2014) User perspectives in public transport timetable optimisation. *Transportation Research Part C: Emerging Technologies* 48(2014):269–284, DOI 10.1016/j.trc.2014.09.005, URL <http://linkinghub.elsevier.com/retrieve/pii/S0968090X14002526>
- Parbo J, Nielsen OA, Prato CG (2016) Passenger Perspectives in Railway Timetabling: A Literature Review. *Transport Reviews* 36(4):500–526, DOI 10.1080/01441647.2015.1113574, URL <https://www.tandfonline.com/doi/full/10.1080/01441647.2015.1113574>
- Paulsen M (2016) Bus and Passenger Simulation within a Combined Agent-Based Multi-Modal Assignment Model. M.sc. thesis, Technical University of Denmark, URL <http://production.datastore.cvt.dk/oafilestore?oid=579743f36bbf232e7000112f&targetid=579743f46bbf232e70001142>
- Poon M, Wong S, Tong C (2004) A dynamic schedule-based model for congested transit networks. *Transportation Research Part B: Methodological* 38(4):343–368, DOI 10.1016/S0191-2615(03)00026-2, URL <http://linkinghub.elsevier.com/retrieve/pii/S0191261503000262>
- Rose G (2006) Mobile Phones as Traffic Probes: Practices, Prospects and Issues. *Transport Reviews* 26(3):275–291, DOI 10.1080/01441640500361108, URL <http://www.tandfonline.com/doi/abs/10.1080/01441640500361108>

- 
- Sato K, Tamura K, Tomii N (2013) A MIP-based timetable rescheduling formulation and algorithm minimizing further inconvenience to passengers. *Journal of Rail Transport Planning and Management* 3(3):38–53, DOI 10.1016/j.jrtpm.2013.10.007, URL <http://dx.doi.org/10.1016/j.jrtpm.2013.10.007>
- Sun H, Wu J, Wu L, Yan X, Gao Z (2016a) Estimating the influence of common disruptions on urban rail transit networks. *Transportation Research Part A: Policy and Practice* 94:62–75, DOI 10.1016/j.tra.2016.09.006, URL <http://linkinghub.elsevier.com/retrieve/pii/S0965856416303913>
- Sun Y, Shi J, Schonfeld PM (2016b) Identifying passenger flow characteristics and evaluating travel time reliability by visualizing AFC data: a case study of Shanghai Metro. *Public Transport* 8(3):341–363, DOI 10.1007/s12469-016-0137-8, URL <http://link.springer.com/10.1007/s12469-016-0137-8>
- Teklu F (2008) A Stochastic Process Approach for Frequency-based Transit Assignment with Strict Capacity Constraints. *Networks and Spatial Economics* 8(2-3):225–240, DOI 10.1007/s11067-007-9046-3, URL <http://link.springer.com/10.1007/s11067-007-9046-3>
- Tong C, Wong S (1999) A stochastic transit assignment model using a dynamic schedule-based network. *Transportation Research Part B: Methodological* 33(2):107–121, DOI 10.1016/S0191-2615(98)00030-7, URL <http://linkinghub.elsevier.com/retrieve/pii/S0191261598000307>
- Xu R, Ye J, Pan H (2014) Method for early warning of heavy passenger flow at transfer station of urban rail transit network under train delay. *Zhongguo Tiedao Kexue/China Railway Science* 35(5), DOI 10.3969/j.issn.1001-4632.2014.05.18
- Zhen Q, Jing S (2016) Train rescheduling model with train delay and passenger impatience time in urban subway network. *Journal of Advanced Transportation* 50(8):1990–2014, DOI 10.1002/atr.1441, URL <http://doi.wiley.com/10.1002/atr.1441>
- Zhu Y, Goverde RMP (2017) Dynamic passenger assignment during disruptions in railway systems. In: 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), IEEE, pp 146–151, DOI 10.1109/MTITS.2017.8005655, URL <http://ieeexplore.ieee.org/document/8005655/>